Clinical Aphasiology, Vol. 23, 1995, pp. 141–154

Prognosis for Improvement During Acute Rehabilitation as Measured by the *Western Aphasia Battery*

Jennifer Horner, Deborah V. Dawson,
Mary Ann Eller, Frances G. Buoyer,
Joy L. Crowder, and Carolyn M. Reus

In their comprehensive review of the aphasia prognosis literature, Rosenbek, LaPointe, and Wertz (1989) found that initial severity is one of the most important prognostic indicators. Pertinent to this important finding, Kertesz (1988) advocated that studies of recovery must "control for" initial severity, and furthermore (Kertesz, 1979) that *amount of improvement* and *ultimate outcome* should be analyzed separately.

Although the discipline of aphasiology knows quite a bit about prognosis, we, and others such as Holland and colleagues (Holland, Greenhouse, Fromm, & Swindell, 1989), continue to revisit the question. We think one reason for this is that researchers study the question of prognosis with groups of patients, whereas in clinical practice each individual patient's prognosis remains elusive. Another reason is that aphasia prognosis studies are plentiful (e.g., Cappa, 1992; Kertesz, 1988; Kertesz & McCabe, 1977; Pashek & Holland, 1988; Vignolo, 1964), but comparative studies across the wide spectrum of cognitive-communication disorders in acute rehabilitation settings are sparse (Duffy & Myers, 1991; Holland et al., 1989). Another reason to revisit the question is the increasing pressure in acute rehabilitation settings to predict and to produce positive treatment outcomes, both in an increasingly diverse group of patients and in an increasingly short period of time.

This study began in 1989. In the first phase, we studied the inter- and intratester reliability of the *Prognosis Profile Worksheet* (PPW) (see Appendix), which we found to be acceptable (Horner, Eller, Dawson, & Buoyer, 1994). The second phase, reported in this paper, assesses the prognostic value of the PPW, as well as other select variables.

The questions posed in this study were: (1) What are the best predictors of outcome as measured by the *Western Aphasia Battery* (Kertesz, 1982) in a representative sample of acute rehabilitation patients? (2) Do patients representing four diagnostic categories differ in posttest or change scores when baseline values are controlled? (3) What is the outcome pattern? (4) Do scores from the PPW add appreciably to prediction of outcome? (5) Do clinical prognostic ratings from the PPW (excellent, good, fair, guarded, poor) add appreciably to prediction of outcome?

METHOD

Subjects

We studied 69 patients from four diagnostic categories (see Table 1). All had aphasia or other cognitive-communication disorders, could participate in standardized testing, and received short-term treatment. Patients with dysarthria or dysphagia only were excluded. Also, patients who underwent one evaluation only were excluded.

All subjects (N = 69) completed the *Western Aphasia Battery* (WAB) (Kertesz, 1982) once prior to treatment and once after treatment (see Table 2). The *Prognosis Profile Worksheet* (PPW) (Horner et al., 1994) was completed by the primary clinician at the time of baseline examination.

Analysis

In our analysis of prognosis, we considered several demographic variables (age, education, sex, and race) and several clinical variables (diagnostic category, days postonset first seen by speech–language pathology, total hospital days, total rehabilitation days including weekends, total treatment hours, and total treatment sessions). In addition, we considered data from the PPW, including subscores for demographic, language, other higher cortical functions, and visual-motor functions. When summed, these subscores yield a total prognosis score (maximum 100 points). Finally, the primary clinician rendered a global prognostic estimate (excellent, good, fair, guarded, poor). Our dependent measure was the WAB. We compared baseline performance with four outcomes: Aphasia Quotient (AQ) posttest scores, AQ change scores, Cortical Quotient (CQ) posttest scores, and CQ change scores.

Table 1. Demographic and Clinical Data for Four Diagnostic Groups (*N* = 69)

	LHS (N = 27)	*RHS* (N = 17)	*BHS* (N = 11)	*NON* (N = 14)
Male; female	19;8	11;6	6;5	12;2
Caucasian; non-Caucasian	19;8	8;9	7;4	9;5
Age (years)	65.2 [12.2]	64.7 [11.4]	60.9 [14.3]	41.9 [15.4]
Education (years)	10.5 [4.0]	11.2 [3.9]	11.3 [3.9]	11.3 [2.6]
Days postonset[a]	22.5 [12.2]	43.1 [78.0]	17.4 [8.2]	38.2 [25.7]
Total hospital days	68.5 [18.8]	65.4 [26.7]	43.5 [12.0]	78.1 [39.2]
Total rehabilitation days	43.3 [13.0]	39.9 [13.9]	26.5 [6.7]	40.6 [15.1]
Total treatment hours	17.6 [7.8]	11.4 [5.8]	11.0 [4.5]	16.0 [8.4]
Total treatment sessions	34.0 [13.0]	23.1 [11.9]	20.0 [9.5]	28.8 [13.4]

Note: Standard deviation provided in brackets. All but one patient was right handed. LHS = left hemisphere stroke; RHS = right hemisphere stroke (14 patients had visual-spatial neglect); BHS = bilateral hemisphere stroke; NON = nonstroke (head injury in 6 patients; neurosurgery in the remaining).
[a]Total days after stroke or injury that patient was seen by speech–language pathology.

The analysis of data was done in two parts. In the first part, we analyzed the relationships among WAB posttest and change scores and WAB baseline scores, demographic variables, and clinical variables. In the second part, we analyzed WAB posttest and change scores relative to baseline scores, PPW scores, and global prognostic estimate. Statistical analyses consisted of Spearmen rank correlations and stepwise regression. We adjusted for multiple comparisons, and we corrected for baseline differences.

RESULTS

First, we compared all the WAB scores across groups using a nonparametric Kruskal-Wallis statistic, and found that aphasic patients having

Table 2. *Western Aphasia Battery* (Kertesz, 1982) Baseline and Outcome Data for Four Diagnostic Groups (N = 69)

	LHS (N = 27)	*RHS* (N = 17)	*BHS* (N = 11)	*NON* (N = 14)
Aphasia Quotient Baseline				
Mean	42.7	89.4	76.2	79.3
SD	31.7	7.9	23.6	22.7
Median	46.6	90.4	84.0	88.9
Aphasia Quotient Posttest				
Mean	54.8	93.8	83.4	89.0
SD	30.5	5.3	20.2	18.6
Median	68.8	95.6	93.0	95.3
Aphasia Quotient Change				
Mean	12.0	4.3	7.2	9.7
SD	10.6	4.4	5.7	6.6
Median	10.0	3.2	8.7	8.1
Cortical Quotient Baseline				
Mean	39.1	79.2	70.6	75.5
SD	26.5	13.9	17.3	20.1
Median	41.8	82.2	71.4	84.6
Cortical Quotient Posttest				
Mean	52.9	84.5	80.6	84.8
SD	27.4	11.2	11.8	15.3
Median	65.9	85.4	83.0	90.9
Cortical Quotient Change				
Mean	13.7	5.3	10.0	9.3
SD	8.3	6.3	9.1	7.3
Median	15.6	5.0	8.9	7.2

Note: LHS = left hemisphere stroke (1 of 27 patients lacked a posttest Cortical Quotient); RHS = right hemisphere stroke; BHS = bilateral hemisphere stroke (1 of 11 patients lacked a Cortical Quotient); NON = nonstroke.

had left hemisphere strokes differed from all other groups, notably because they had a much wider range of scores (p = .019 for AQ change; p = .01 for CQ change; p = .0001 for AQ and CQ baseline scores; and p = .0001 for AQ and CQ posttest scores).

Analysis 1

Table 3 shows correlations obtained between our four WAB "outcome" scores and the baseline AQ and CQ scores as well as the clinical vari-

Table 3. Spearman Rank Correlations Between *Western Aphasia Battery* (WAB) Scores and Demographic Variables (N = 69)

WAB Outcome Variables	*Aphasia Quotient Baseline*	*Cortical Quotient Baseline*	*Time Post-Onset*	*Total Hospital Days*	*Total Rehab Days*	*Total Treatment Hours*	*Total Treatment Sessions*
Aphasia Quotient Posttest	0.916 *	0.877 *	0.033	–0.270	–0.343	–0.540 *	–0.518 *
Cortical Quotient Posttest	0.857 *	0.931 *	–0.013	–0.338	–0.323	–0.439 *	–0.416 *
Change in Aphasia Quotient	–0.662 *	–0.588 *	–0.046	0.195	0.239	0.339	0.308
Change in Cortical Quotient	–0.664 *	–0.634 *	0.010	0.212	0.292	0.422 *	0.388

*P-values less than .0014 are significant after adjustment for multiple comparisons, using Bonferroni adjustment with an overall significance level of .05.

ables of interest. All demographic variables—age, sex, race, and education—failed to correlate significantly with our outcome variables. Thus, these were excluded from further analyses. Next, select outcome variables correlated significantly with treatment hours and sessions. AQ and CQ *posttest* scores correlated significantly with AQ and CQ baseline scores. Similarly, AQ and CQ *change scores* correlated significantly with AQ and CQ baseline scores.

Figures 1 and 2 show the pattern of outcome, using CQ data for all 69 subjects. Figure 1 illustrates that the relationship between CQ baseline

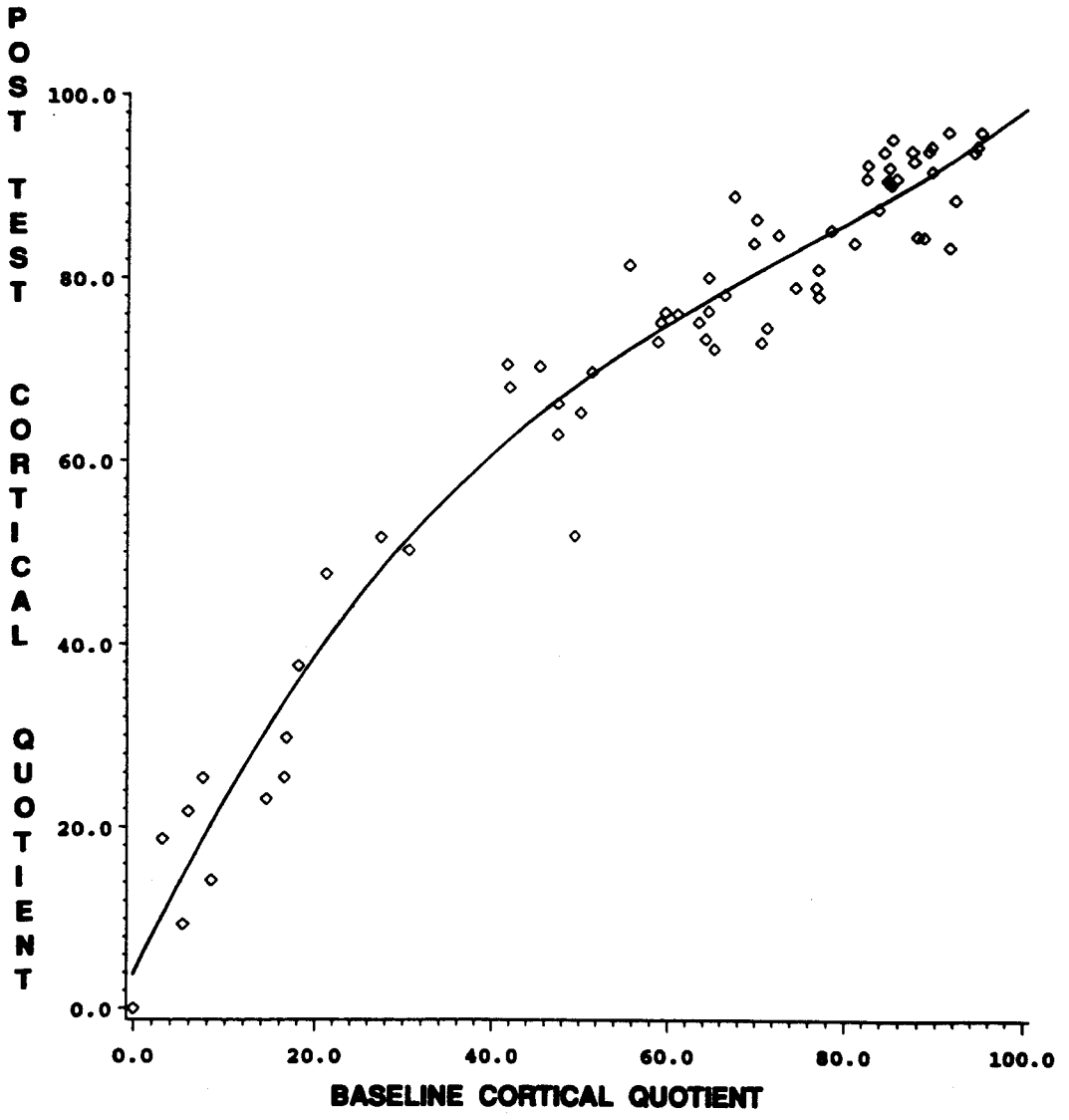

Figure 1. The relationship between posttest cortical quotients relative to baseline cortical quotients (N = 69) is shown to be almost linear.

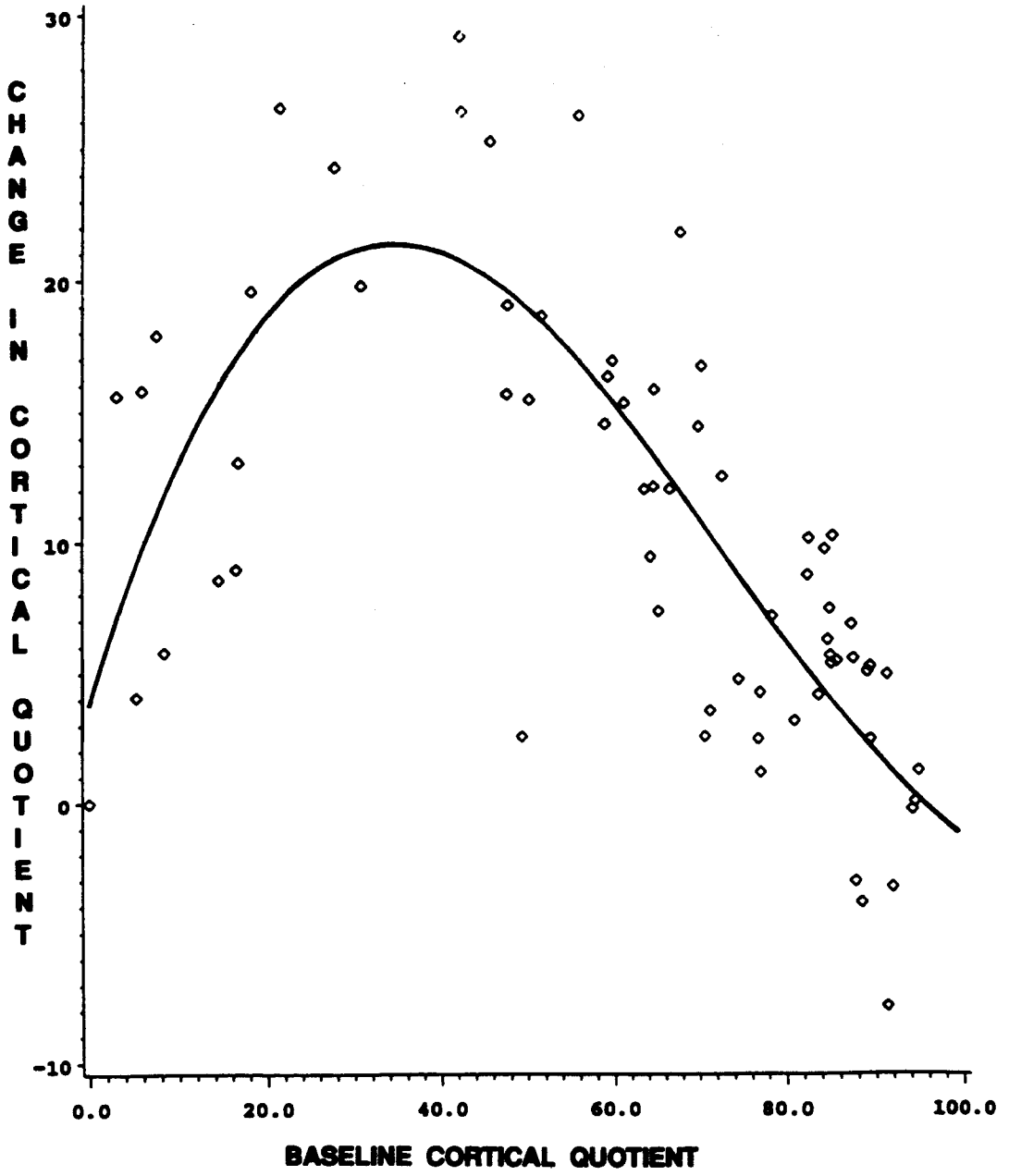

Figure 2. The relationship between cortical quotient change relative to baseline cortical quotients (N = 69) is shown to be close to U-shaped.

and *posttest* CQ performances was roughly linear. Figure 2 shows that the relationship between baseline CQ and *change* in CQ is roughly U-shaped. Patients performing at the lowest and highest ends of the CQ range changed least of all. Patients in the middle range demonstrated relatively greater gains.

Stepwise regression procedures were used to develop a model for each of the four outcomes, using a cubic polynomial to model the relationship with baseline values. The results of this analysis were: (1) the model based upon baseline AQ alone explained 95.8% of the variability in AQ posttest scores ($p < .001$; $R^2 = 95.8\%$); (2) the model based upon baseline AQ alone explained 52.4% of the variability in AQ change scores ($p < .001$, $R^2 = 52.4\%$); (3) the model based upon CQ baseline alone explained 95.7% of the variability in CQ change scores ($p < .001$, $R^2 = 95.7\%$); and (4) the model based upon CQ baseline alone explained 61.5% of the variability in CQ change scores ($p < .001$, $R^2 = 61.5\%$). Only one additional variable—number of treatment sessions—was retained in the regression models pertaining to only two outcomes, CQ posttest and CQ change scores. In both instances, the predictive value increased only modestly ($R^2 = 95.7\%$ increased to $R^2 = 95.9\%$; and $R^2 = 61.5\%$ increased to $R^2 = 63.2\%$, respectively). It is noteworthy that diagnostic category was not retained in the model, despite the intergroup differences noted earlier. Once an adjustment for baseline values was made, no additional information was provided by diagnostic category. We found no evidence that a model specifying separate predictive questions for each diagnostic group provided a significantly better "fit" than the single model applied across all diagnostic groups ($p > .20$ for all outcomes). Therefore, a separate statistical equation was not needed for the four diagnostic groups.

Analysis 2

Table 4 summarizes the correlations of WAB outcome scores with the PPW subscores, the PPW total scores, and the global ratings. In this correlational analysis, the PPW total scores correlated significantly with AQ and CQ posttest scores, *and* with AQ and CQ change scores. In contrast, the PPW global prognosis ratings correlated significantly with AQ and CQ posttest scores, but failed to correlate significantly with AQ and CQ change scores.

When we subjected these PPW data to a stepwise regression with baseline AQ and CQ values to assess whether these components added to the predictive ability of the models, we found that none of the PPW scores added significantly to the prediction of AQ and CQ outcomes. Again (as in the previous analysis), AQ and CQ baseline scores explained over 95% of the variability in AQ and CQ *posttest scores*, respectively. AQ and CQ baseline scores explained over 50% of the variability in AQ and CQ *change scores*, respectively.

Table 4. Spearman Rank Correlations Between *Western Aphasia Battery* (WAB) Scores and Scores from the *Prognosis Profile Worksheet (PPW)* (*N* = 69)

WAB Outcome Variables	*PPW Demographics Score*	*PPW Language Subscore*	*PPW Higher Cortical Abilities Subscore*	*PPW Visual-Motor Abilities Subscore*	*PPW Total Prognosis Score (max = 100)*	*PPW Global Prognosis Rating*
Aphasia Quotient Posttest	0.05	0.803 *	0.555 *	0.68 *	0.712 *	0.444 *
Cortical Quotient Posttest	–0.052	0.877 *	0.704 *	0.678 *	0.79 *	0.545 *
Aphasia Quotient Change	0.203	–0.492 *	–0.475 *	–0.362 *	–0.477 *	–0.241
Cortical Quotient Change	0.002	–0.542 *	–0.439 *	0.27 *	–0.501 *	–0.252

**P*-values less than .0021 are significant after adjustment for multiple comparisons, using Bonferroni adjustment with an overall significance level of 5%.

Summary

First, we found that the AQ and CQ baseline scores were the best predictors of cognitive-communication outcomes as measured in this study. Second, patients representing four diagnostic categories did not differ in posttest or change scores when baseline values were controlled. Third, the pattern of outcome was roughly linear when comparing baseline and posttest scores, and was roughly U-shaped when comparing baseline and change scores. Fourth, scores from the PPW did not add appreciably to prediction of outcome (when baseline was taken into account). Fifth, global prognosis ratings also did not add appreciably to prediction of outcome (when baseline was taken into account).

DISCUSSION

We recognize a number of limitations to this study. First, our test battery was not diagnosis (etiology) specific. In the future, we plan to be more selective in our choice of tests. Second, not all patients with cognitive-communication impairment at Duke Rehabilitation Center who were potentially eligible for an acute rehabilitation prognosis study were testable using the WAB. Third, we had a variable number of subjects across diagnostic groups, and variable levels of severity. (This was not a cohort study as advocated by Rosenbek et al., 1989.) Fourth, the duration of treatment was short, sometimes unexpectedly so, and the amount of treatment was variable. Fifth, treatment rendered was individualized, not specified. Sixth, lesion variables, undoubtedly of great potential value to the question of prognosis, were not included in this study.

We address clinical implications in the form of questions:

1. Can we conclude from this study that speech–language treatment has little effect? No. Our change scores were small but, we hope, not insignificant. We emphasize this was not a treatment study, but rather a study of *change in performance* on the WAB in a realistic setting.
2. Do our results suggest that we should treat only moderately impaired patients, that is, those in the middle of the U-shaped curve? No, because as clinicians, we know that often small changes in severely or mildly impaired patients *are clinically significant*. Rather, our data suggest that the WAB probably does not capture as well as we would like the small, pos-

sibly significant changes at the lower and upper ranges of severity.

3. Should we study only long-term outcome, not short-term? Again, no. In spite of the short hospital stays of our patients, or perhaps because of them, we are compelled to sharpen our prognostication skills. Our quandary is how best to do this.
4. Should we discard the *Western Aphasia Battery?* No, but we think we should use the WAB for left hemisphere–damaged aphasic patients and, for other patients, rely more heavily on the etiology-specific tests now available.
5. Should we discard the *Prognosis Profile Worksheet*? Frankly, we are not sure. Some of the worksheet scores correlated highly with the AQ and CQ outcomes, but did not add appreciably to the prediction above and beyond that provided by AQ and CQ baseline scores.
6. Should we discard the global prognostic ratings or estimates? The short answer, based on the results of this study, is, perhaps, yes! The long answer is no, as follows. Horner and colleagues (1994) began with a quote from Wertz (1978), who said of prognostication, "The task is to make a prospective statement for each patient individually" (p. 26). Both Wertz (1978) and Darley (1982) advised clinicians to ask "prognosis . . . for what . . . for whom . . . and . . . at what point in time?" Of all of the components of this question, the WHAT part bothers us the most. *What* do we mean by WHAT? Or to paraphrase Joe Duffy's question at the Twenty-Second Annual Clinical Aphasiology Conference: "WHAT are you prognosticating?"

In an attempt to answer that question, we suggest to you that clinicians prognosticate *about anticipated change* and treat patients *in anticipation of change.* The challenge of prognostication is to go beyond anticipation and to make accurate, explicit judgments about the *likelihood* of change, about the *magnitude* of change, and about the clinical *relevance* of these changes. Also, the central challenge of prognostication is to focus these explicit judgments on *specific neurobehavioral domains;* that is, it seems reasonable to us to use global prognostic ratings (e.g., excellent, good, fair, guarded, poor) as long as the behaviors about which we prognosticate are *specific* and *measurable.*

Finally, and perhaps most important, we think the data in this report are telling us that we should study the prognosis of individuals, not of groups. That accurate prognostic statements are important is undeniable. We hope that aphasiologists continue to revisit the research ques-

tions surrounding prognosis until we fully understand "for what . . . for whom . . . and at what point in time" (Darley, 1982; Wertz, 1978).

REFERENCES

Cappa, S. F. (1992). Recovery of aphasia [Special Issue]. *Aphasiology, 6*(4).

Darley, F. L. (1982). *Aphasia* (pp. 110–143). Philadelphia: Saunders.

Duffy, J. R., & Myers, P. S. (1991). Group comparisons across neurologic communication disorders: Some methodological issues. In T. E. Prescott (Ed.), *Clinical aphasiology* (Vol. 19, pp. 1–14). Austin TX: PRO-ED.

Holland, A. L., Greenhouse, J. B., Fromm, D., & Swindell, C. S. (1989). Predictors of language restitution following stroke: A multivariate analysis. *Journal of Speech and Hearing Research, 32*, 232–238.

Horner, J., Eller, M. A., Dawson, D. V., & Buoyer, F. G. (1994). Duke Prognosis Profile Worksheet: Inter- and intra-observer reliability. In M. L. Lemme (Ed.), *Clinical aphasiology* (Vol. 22, pp. 107–118). Austin, TX: PRO-ED

Horner, J., & Rothi, L. J. (1984). Prognosis for recovery from aphasia: Estimating extent of improvement. *Communique* (North Carolina Speech-Language-Hearing Association Journal) Research Edition: pp. 20–29.

Kertesz, A. (1979). *Aphasia and associated disorders: Taxonomy, localization, and recovery*. New York: Grune & Stratton.

Kertesz, A. (1982). *Western Aphasia Battery*. New York: Grune & Stratton.

Kertesz, A. (1988). What do we learn from recovery from aphasia? In S. G. Waxman (Ed.), *Advances in neurology: Functional recovery in neurological disease* (Vol. 47, pp. 277–292). New York: Raven Press.

Kertesz, A., & McCabe, P. (1977). Recovery patterns and prognosis in aphasia. *Brain, 100*, 1–18.

Pashek, G. V., & Holland, A. L. (1988). Evolution of aphasia in the first year post-onset. *Cortex, 24*, 411–423.

Rosenbek, J. C., LaPointe, L. L., & Wertz, R. T. (1989). *Aphasia: A clinical approach* (pp. 90–99). Boston: Little, Brown/College-Hill.

Vignolo, L. A. (1964). Evolution of aphasia and language rehabilitation: A retrospective exploratory study. *Cortex, 1*, 344–367.

Wertz, R. T. (1978). Neuropathologies of speech and language: An introduction to patient management. In D. F. Johns (Ed.), *Clinical management of neurogenic communicative disorders* (pp. 1–101). Boston: Little, Brown.

APPENDIX:
DUKE PROGNOSIS PROFILE WORKSHEET

ESTIMATING POTENTIAL FOR EXTENT OF IMPROVEMENT

Patient:
Date of Onset:

Date of Evaluation:
Clinician:

ETIOLOGY

STROKE
- ..... Ischemic
  - ..... Embolic
  - ..... Thrombotic
- ..... Hemorrhagic
  - ..... Intracerebral (ICH)
  - ..... Subarachnoid (SAH)
    - ..... Aneurysm
    - ..... Arteriovenous malformation (AVM)

TRAUMA
- ..... Closed head injury (CHI)
  - ..... Concussion
  - ..... Contusion
  - ..... Subdural hematoma (SDH)
- ..... Intracerebral hemorrhage (ICH)
  - ..... Coup ..... Contre coup
  - ..... Open head injury (OHI)/with skull fracture
- ..... Concussion
  - ..... Contusion
  - ..... SDH
  - ..... ICH
  - ..... Coup ..... Contre-coup

TUMOR
- ..... Glioblastoma Multiforme
- ..... Astrocytoma
- ..... Meningioma
- ..... Metastatic
- ..... Other:

..... Loss of consciousness – or coma RLA I, II,III – and duration:

Neuroradiology:

..... Head CT Scan ..... Head MRI ..... Angiography ..... PET or SPECT ..... Other

Aphasia Profile (W.A.B. - A.Q.) [If applicable]

..... Anomic
..... Conduction
..... Wernicke
..... Transcortical Sensory
..... Broca
..... Transcortical motor
..... Global
..... Mixed transcortical
..... Unclassifiable; NOT aphasic; normal
..... Unclassifiable; NOT aphasic; ABnormal

ADDITIONAL VARIABLES:

Race: Caucasian Black Hispanic Other:

Occupation (now or in past):

Number of TOTAL hospital days (i.e., from acute hospital admission through discharge from rehabilitation) :

RATING:		3	2	1
HANDEDNESS:		..... Left	..... Right	..... Ambidextrous
SEX:		..... Male	..... Female	
AGE:		..... ≤ 40	 41-69	..... ≥70
EDUCATION:		..... >12	 9-12	..... <9
LESION TYPE:	Stroke	..... Hemmorhagic	..... Ischemic	..... Hemorrhagic infarct
	Trauma	..... Focal	..... Diffuse	..... Focal and diffuse
	Tumor	..... Slow growing	..... Rapid growing	
LESION NUMBER:		..... Single	..... Multiple	..... Multiple with atrophy
LESION SIDE:		..... Right	..... Left	..... Bilateral
NEUROLOGIC SIGNS:		..... No weakness	..... L or R arm/leg	..... Bilateral arm/leg
VISUAL FIELD DEFECT:		..... No	..... Yes	..... (Neglect precludes unequivocal test)
MONTHS POST ONSET:		 0-1	 1-6	..... > 6

{ /30}

RATING:	5	4	3	2	1	0 = Not testable
	Normal	Mild	Moderate	Severe	Profound	
LANGUAGE PERFORMANCE:						
Auditory	.....	.....	.....	.....	.....	{ /5 }
Oral-verbal	.....	.....	.....	.....	.....	{ /5 }
Reading	.....	.....	.....	.....	.....	{ /5 }
Writing	.....	.....	.....	.....	.....	{ /5 }
Subscore:						{ /20}
OTHER HIGHER CORTICAL FUNCTIONS:						
Construction (Copying, Drawing, Blocks)	.....	.....	.....	.....	.....	{ /5 }
Calculations	.....	.....	.....	.....	.....	{ /5 }
Abstract reasoning	.....	.....	.....	.....	.....	{ /5 }
Verbal memory	.....	.....	.....	.....	.....	{ /5 }
Visual memory	.....	.....	.....	.....	.....	{ /5 }
Subscore:						{ /25}
VISUAL - MOTOR FUNCTIONS:						
Praxis: Buccofacial	.....	.....	.....	.....	.....	{ /5 }
Praxis: Limb	.....	.....	.....	.....	.....	{ /5 }
Praxis: Verbal	.....	.....	.....	.....	.....	{ /5 }
Dysarthria	.....	.....	.....	.....	.....	{ /5 }
Visual neglect	.....	.....	.....	.....	.....	{ /5 }
Subscore:						{ /25}
TOTAL:						{ /100}

OVERALL PROGNOSTIC ESTIMATE FOR EXTENT OF IMPROVEMENT IN APHASIA/COGNITION/COMMUNICATION

..... Excellent ..... Good ..... Fair ..... Guarded ..... Poor

From "Duke *Prognosis Profile Worksheet:* Inter- and Intraobserver Reliability" by J. Horner, M. A. Eller, D. V. Dawson, & F G. Buoyer, in *Clinical Aphasiology* (Vol. 22, pp. 117–118) by M. L. Lemme (Ed.), 1994, Austin, TX: PRO-ED. Reprinted with permission. (Based on Horner & Rothi, 1984.)